wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:28 AM GMT

PDB ID	:	2X6I
Title	:	THE CRYSTAL STRUCTURE OF THE DROSOPHILA CLASS III PI3-KINASE VPS34 IN COMPLEX WITH PIK-90
Authors	:	Miller, S.; Tavshanjian, B.; Oleksy, A.; Perisic, O.; Houseman, B.T.; Shokat, K.M.; Williams, R.L.
Deposited on	:	2010-02-17
Resolution	:	3.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

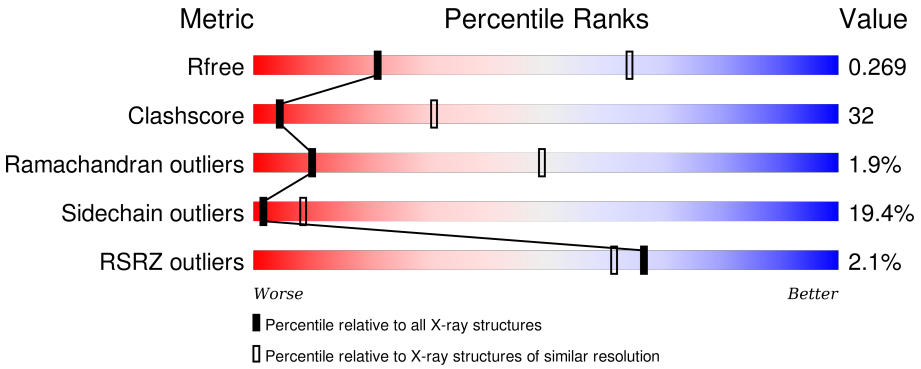
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	696	% 37% 30% 11% • 22%
1	B	696	2% 38% 31% 9% 22%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	090	B	1950	-	-	X	-

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8976 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

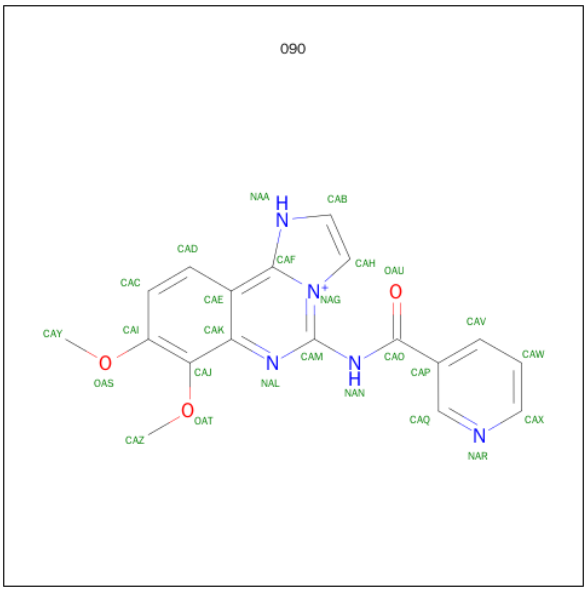
- Molecule 1 is a protein called PHOSPHOTIDYLINOSITOL 3 KINASE 59F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	546	Total	C	N	O	S	0	0	0
			4468	2888	761	792	27			
1	B	544	Total	C	N	O	S	0	0	0
			4456	2881	760	788	27			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	254	GLY	-	EXPRESSION TAG	UNP Q9W1M7
A	255	SER	-	EXPRESSION TAG	UNP Q9W1M7
A	256	HIS	-	EXPRESSION TAG	UNP Q9W1M7
A	257	MET	-	EXPRESSION TAG	UNP Q9W1M7
A	455	ALA	GLY	ENGINEERED MUTATION	UNP Q9W1M7
B	254	GLY	-	EXPRESSION TAG	UNP Q9W1M7
B	255	SER	-	EXPRESSION TAG	UNP Q9W1M7
B	256	HIS	-	EXPRESSION TAG	UNP Q9W1M7
B	257	MET	-	EXPRESSION TAG	UNP Q9W1M7
B	455	ALA	GLY	ENGINEERED MUTATION	UNP Q9W1M7

- Molecule 2 is N-(2,3-DIHYDRO-7,8-DIMETHOXYIMIDAZO[1,2-C] QUINAZOLIN-5-YL) NICOTINAMIDE (three-letter code: 090) (formula: $C_{18}H_{16}N_5O_3$).



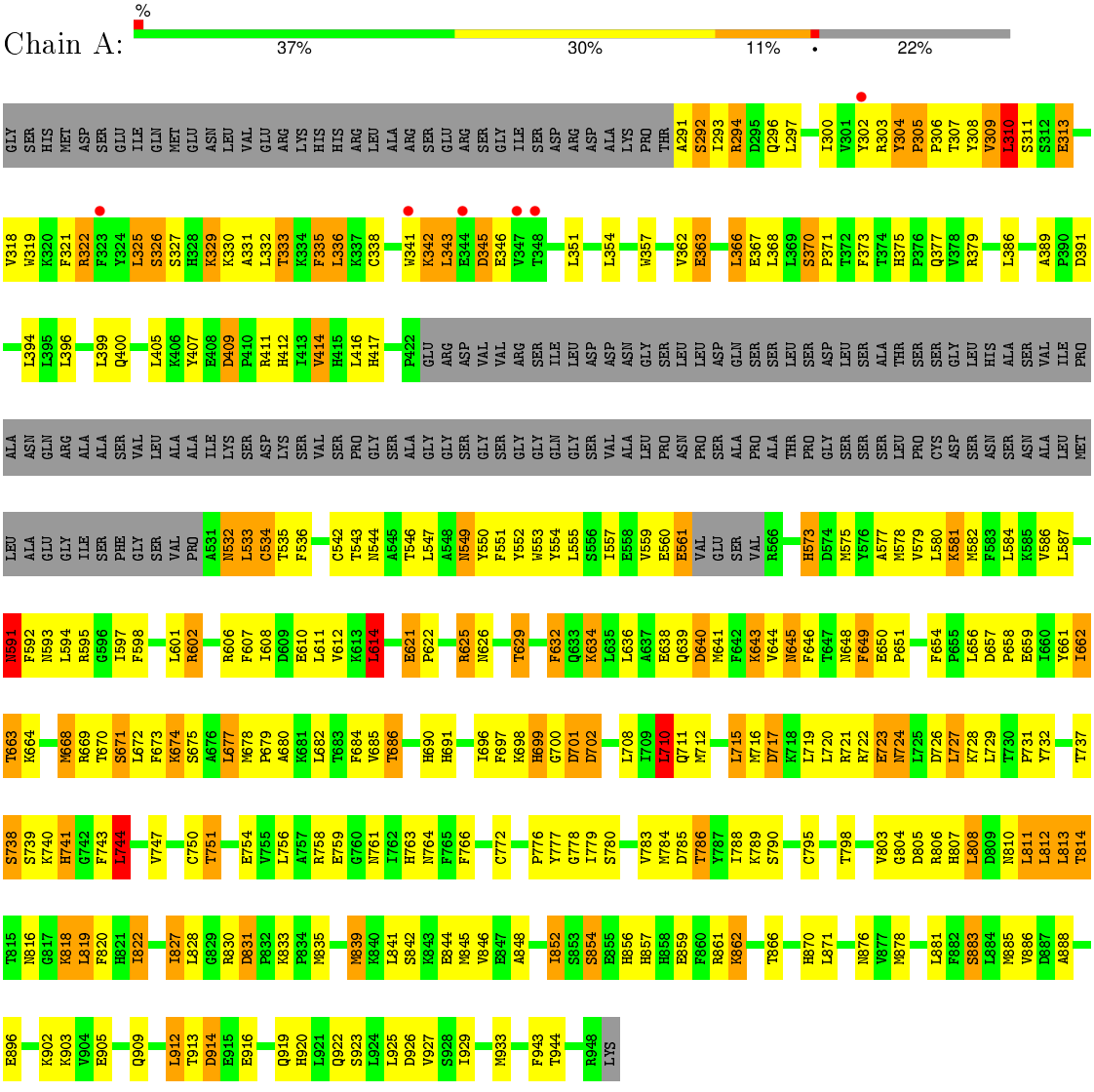
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			26	18	5	3		
2	B	1	Total	C	N	O	0	0
			26	18	5	3		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

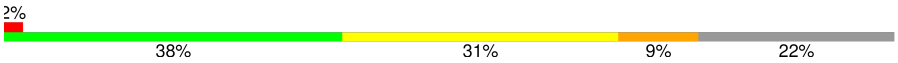
- Molecule 1: PHOSPHOTIDYLINOSITOL 3 KINASE 59F

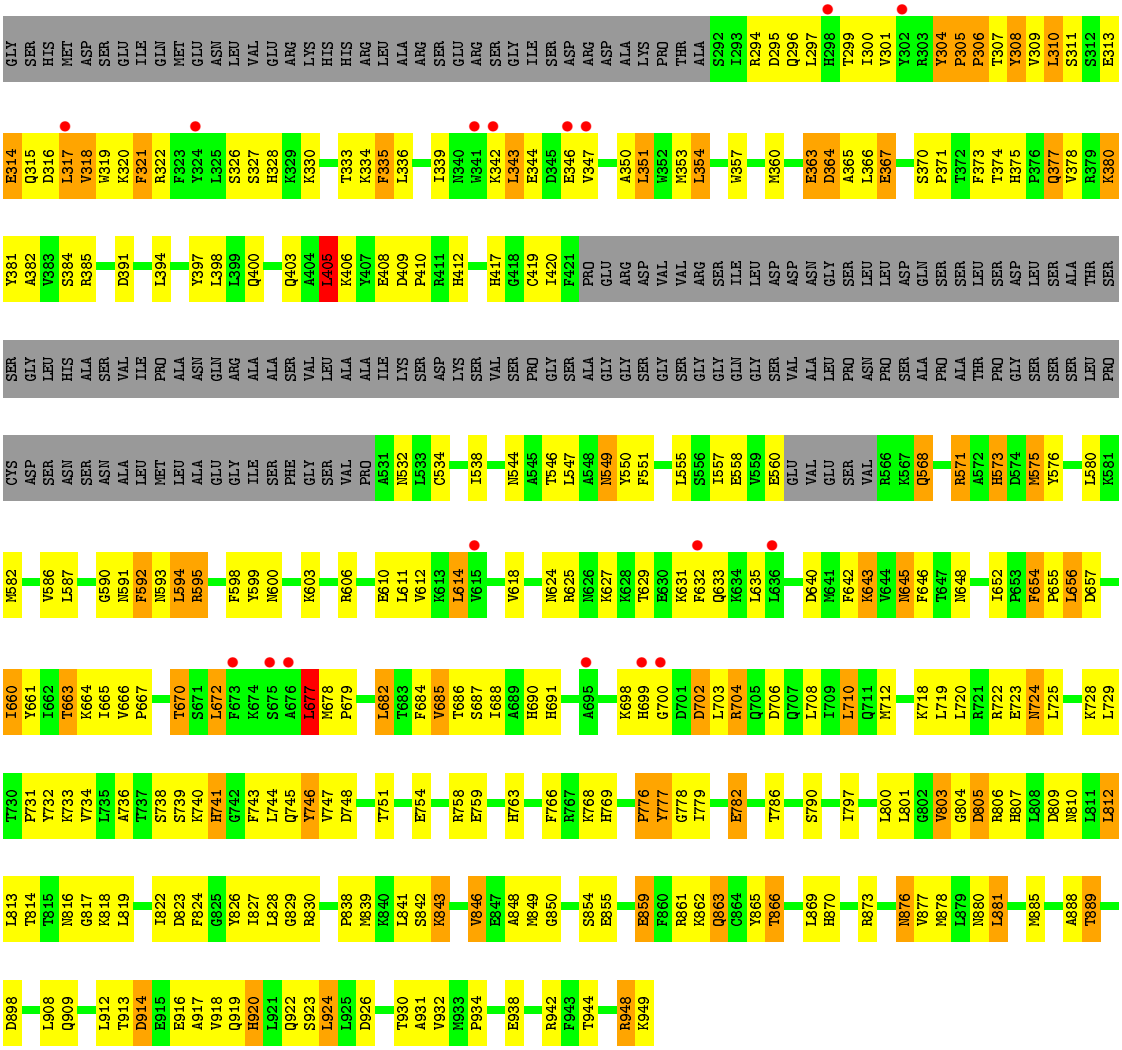
Chain A:



- Molecule 1: PHOSPHOTIDYLINOSITOL 3 KINASE 59F

Chain B:




GLY SER HIS MET ASP SER GLU ILE GLN MET GLU ASN LEU VAL GLU ARG LYS HIS HIS ARG LEU ALA ARG SER GLU ARG SER GLY ILE SER ASP ARG ASP ALA LYS PRO THR ALA S292 I293 R294 D295 Q296 L297 H298 T299 I300 V301 Y302 R303 Y304 P305 P306 T307 Y308 V309 L310 S311 S312 E313
E314 Q315 D316 L317 V318 W319 K320 F321 R322 F323 Y324 L325 S326 S327 H328 K329 K330 T333 K334 F335 L336 I339 N340 W341 K342 L343 E344 D345 E346 V347 A350 L351 W352 M353 L354 W357 M360 E363 D364 A365 L366 E367 S370 P371 T372 F373 T374 H375 P376 Q377 V378 R379 K380
Y381 A382 V383 S384 R385 D391 L394 Y397 L398 L399 Q400 Q403 A404 L405 K406 Y407 E408 D409 P410 R411 H412 H417 G418 C419 I420 F421 PRO GLU ARG ASP VAL VAL ARG SER ILE LEU LEU ASP ASP ASN GLY SER LEU LEU LEU ASP GLN SER SER LEU SER ASP LEU SER ALA THR SER
SER GLY LEU HIS ALA SER VAL ILE PRO ALA ASN GLN ARG ALA ALA ALA SER VAL LEU ALA ALA ILE LYS SER ASP LYS SER VAL SER PRO GLY SER ALA GLY GLY SER GLY SER GLY GLY GLN GLY SER VAL ALA LEU PRO ASN PRO SER ALA PRO ALA ALA THR PRO GLY SER SER SER LEU PRO
CYS ASP SER ASN SER ASN ALA LEU MET LEU ALA GLU GLY ILE SER PHE GLY SER VAL PRO A531 N532 L533 C534 I538 N544 A545 T546 L547 A548 N549 Y550 F551 L555 S556 I557 E558 V559 E560 GLU VAL GLU SER VAL R566 K567 Q568 R571 A572 H573 D574 M575 Y576 L580 K581
M582 V586 L587 G590 N591 F592 N593 L594 R595 F598 Y599 N600 K603 R606 E610 L611 V612 K613 L614 V615 V618 N624 R625 N626 K627 K628 T629 E630 K631 F632 Q633 K634 L635 L636 D640 M641 F642 K643 V644 N645 F646 T647 N648 I652 F653 F654 P655 L656 D657
I660 Y661 I662 T663 K664 I665 V666 P667 T670 S671 L672 F673 K674 S675 A676 L677 M678 P679 L682 T683 F684 V685 T686 S687 I688 A689 H690 H691 A695 K698 H699 G700 D701 D702 L703 R704 Q705 D706 Q707 L708 I709 L710 Q711 M712 K718 L719 L720 R721 R722 E723 N724 L725 K728 L729
T730 F731 Y732 K733 V734 L735 A736 T737 S738 S739 K740 H741 G742 F743 L744 Q745 Y746 V747 D748 T751 E754 R758 E759 H763 F766 R767 K768 H769 P776 Y777 G778 I779 E782 T786 S790 I797 L800 L801 G802 V803 G804 D805 R806 H807 L808 D809 N810 L811 L812
L813 T814 T815 N816 G817 K818 L819 I822 D823 F824 G825 Y826 I827 L828 G829 R830 P838 M839 K840 L841 S842 K843 V846 E847 A848 M849 G850 S854 E855 E859 F860 R861 K862 Q863 C864 Y865 T866 L869 H870 R873 N876 V877 M878 L879 N880 L881 M885 A888 T889
D898 L908 Q909 L912 T913 D914 E915 E916 A917 V918 Q919 H920 L921 Q922 S923 L924 L925 D926 T930 A931 V932 M933 P934 E938 R942 F943 T944 R948 K949

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.47Å 156.22Å 244.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.83 – 3.40 72.83 – 3.40	Depositor EDS
% Data completeness (in resolution range)	97.6 (72.83-3.40) 97.7 (72.83-3.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 3.41Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.217 , 0.279 0.211 , 0.269	Depositor DCC
R_{free} test set	1426 reflections (5.16%)	DCC
Wilson B-factor (Å^2)	96.9	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 61.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 29052 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8976	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 090

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.79	2/4577 (0.0%)	0.91	7/6197 (0.1%)
1	B	0.62	0/4564	0.75	4/6177 (0.1%)
All	All	0.71	2/9141 (0.0%)	0.83	11/12374 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	292	SER	CB-OG	18.05	1.65	1.42
1	A	795	CYS	CB-SG	-5.53	1.72	1.81

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	710	LEU	CA-CB-CG	-7.40	98.28	115.30
1	B	710	LEU	CA-CB-CG	-7.29	98.54	115.30
1	B	677	LEU	CA-CB-CG	7.13	131.71	115.30
1	B	335	PHE	CB-CG-CD2	6.36	125.25	120.80
1	A	854	SER	N-CA-C	-5.92	95.00	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	591	ASN	Mainchain
1	A	663	THR	Mainchain
1	A	680	ALA	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4468	0	4503	319	0
1	B	4456	0	4498	274	0
2	A	26	0	16	3	0
2	B	26	0	16	10	0
All	All	8976	0	9033	585	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 32.

The worst 5 of 585 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:292:SER:CB	1:A:292:SER:OG	1.65	1.42
1:B:595:ARG:HG2	1:B:595:ARG:HH11	1.09	1.12
1:B:311:SER:HB2	1:B:314:GLU:HG3	1.24	1.11
1:A:677:LEU:HD11	1:A:700:GLY:H	1.18	1.05
1:B:311:SER:HB2	1:B:314:GLU:CG	1.87	1.05

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was

analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	540/696 (78%)	504 (93%)	26 (5%)	10 (2%)	10 49
1	B	538/696 (77%)	503 (94%)	24 (4%)	11 (2%)	9 48
All	All	1078/1392 (77%)	1007 (93%)	50 (5%)	21 (2%)	10 49

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	343	LEU
1	B	592	PHE
1	B	778	GLY
1	A	305	PRO
1	A	310	LEU

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	491/612 (80%)	389 (79%)	102 (21%)	1 6
1	B	490/612 (80%)	402 (82%)	88 (18%)	2 11
All	All	981/1224 (80%)	791 (81%)	190 (19%)	2 8

5 of 190 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	813	LEU
1	B	313	GLU
1	B	855	GLU
1	A	822	ILE
1	A	883	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	340	ASN
1	B	532	ASN
1	B	857	HIS
1	B	349	GLN
1	B	549	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	090	A	1949	-	26,29,29	4.63	14 (53%)	30,41,41	1.76	8 (26%)
2	090	B	1950	-	26,29,29	4.85	15 (57%)	30,41,41	1.70	5 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	090	A	1949	-	-	0/12/12/12	0/4/4/4
2	090	B	1950	-	-	0/12/12/12	0/4/4/4

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1949	090	CAJ-CAK	-4.77	1.36	1.42
2	B	1950	090	CAJ-CAK	-3.17	1.38	1.42
2	B	1950	090	CAV-CAP	-2.43	1.35	1.39
2	A	1949	090	CAX-NAR	2.09	1.40	1.33
2	B	1950	090	CAX-NAR	2.15	1.40	1.33

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1950	090	CAB-CAH-NAG	-5.98	102.27	106.97
2	A	1949	090	CAB-CAH-NAG	-4.13	103.72	106.97
2	A	1949	090	OAS-CAI-CAC	-3.44	118.56	124.35
2	B	1950	090	CAE-CAK-NAL	-2.92	119.67	123.42
2	A	1949	090	CAD-CAE-CAF	-2.49	119.04	122.61

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1949	090	3	0
2	B	1950	090	10	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	546/696 (78%)	0.11	6 (1%)	82	77	36, 59, 97, 114	0
1	B	544/696 (78%)	0.25	17 (3%)	52	48	47, 70, 100, 110	0
All	All	1090/1392 (78%)	0.18	23 (2%)	67	61	36, 65, 99, 114	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	302	TYR	6.1
1	B	341	TRP	4.6
1	B	302	TYR	4.2
1	B	324	TYR	4.1
1	A	341	TRP	3.4

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	090	A	1949	26/26	0.91	0.36	1.68	69,77,87,88	0
2	090	B	1950	26/26	0.94	0.30	-0.06	75,80,95,96	0

6.5 Other polymers (i)

There are no such residues in this entry.